# An Introduction to Campus Grids
# 19-Apr-2010

Keith Chadwick

&

Steve Timm

# Outline

- Definition of a Campus Grid
- Why do Campus Grids
- Drawbacks of Campus Grids
- Examples of Campus Grids
  - GLOW
  - Purdue
  - University of California
  - Nebraska
- FermiGrid
  - Pre-Grid Situation
  - Architecture
  - Metrics
- Evolution & Other Considerations
- Cloud Computing
- Additional Resources
- Conclusions

# Definition

- A Campus Grid is a distributed collection of [compute and storage] resources, provisioned by one or more stakeholders, that can be seamlessly accessed through one or more [Grid] portals.

# Why Do Campus Grids ?

- Improve utilization of (existing) resources – don't purchase resources when they are not needed.
  - Cost savings.
- Provide common administrative framework and user experience.
  - Cost savings.
- Buy resources (clusters) in "bulk" @ lower costs.
  - Cost savings.
- Lower maintenance costs.
  - Cost savings.
- Unified user interface will reduce the amount of user training required to make effective use of the resources.
  - Cost savings.

# What are the drawbacks ?

- Additional centralized infrastructure to provision and support.
  - Additional costs.
  - Can be provisioned incrementally to manage buy-in costs.
  - Virtual machines can be used to lower buy-in costs.

- Can make problem diagnosis somewhat more complicated.
  - Correlation of multiple logs across administrative boundaries.
  - A central log repository is one mechanism to manage this.

- Not appropriate for all workloads.
  - Don't want campus financials running on the same resources as research.

- Have to learn (and teach the user community) how to route jobs to the appropriate resources.
  - Trivially parallel jobs require different resources than MPI jobs.
  - I/O intensive jobs require different resources than compute intensive jobs.

- Limited stakeholder buy-in may lead to a campus grid that's less interoperable than you might like.

# GLOW

- Single Globus Gatekeeper (GLOW)
- Large central cluster funded by grant.
- Multiple department based clusters all running Condor.
- Departments have priority [preemptive] access to their clusters.
- Clusters interchange workloads using Condor "flocking".
- Approximately 1/3 of jobs are opportunistic.

# Purdue

- Single Gatekeeper (Purdue-Steele)
- Centrally managed "Steele" cluster.
- ??? Nodes, ??? Slots
- Departments purchase "slots" on the cluster.
- Primary batch scheduler is PBS for purchased slots.
- Secondary batch scheduler is Condor for opportunistic computing.
- Condor is configured to only run jobs when PBS is not running a job on the node.

# University of California

- Multiple campuses.
- Each campus has a local campus Grid portal.
- Overall Grid portal in addition.
- Access is Web portal based.

# Nebraska

- 3 Campuses across Nebraska.
- Being commissioned now.

# Fermilab – Pre Grid

- Multiple "siloed" clusters, each dedicated to a particular stakeholder:
  - CDF – 2 clusters, ~2,000 slots
  - D0 – 2 clusters, ~2,000 slots
  - CMS – 1 cluster, ~4,000 slots
  - GP – 1 cluster, ~500 slots

- Difficult to share:
  - When a stakeholder needed more resources, or did not need all of their currently allocated resources, it was extremely difficult to move jobs or resources to match the demand.

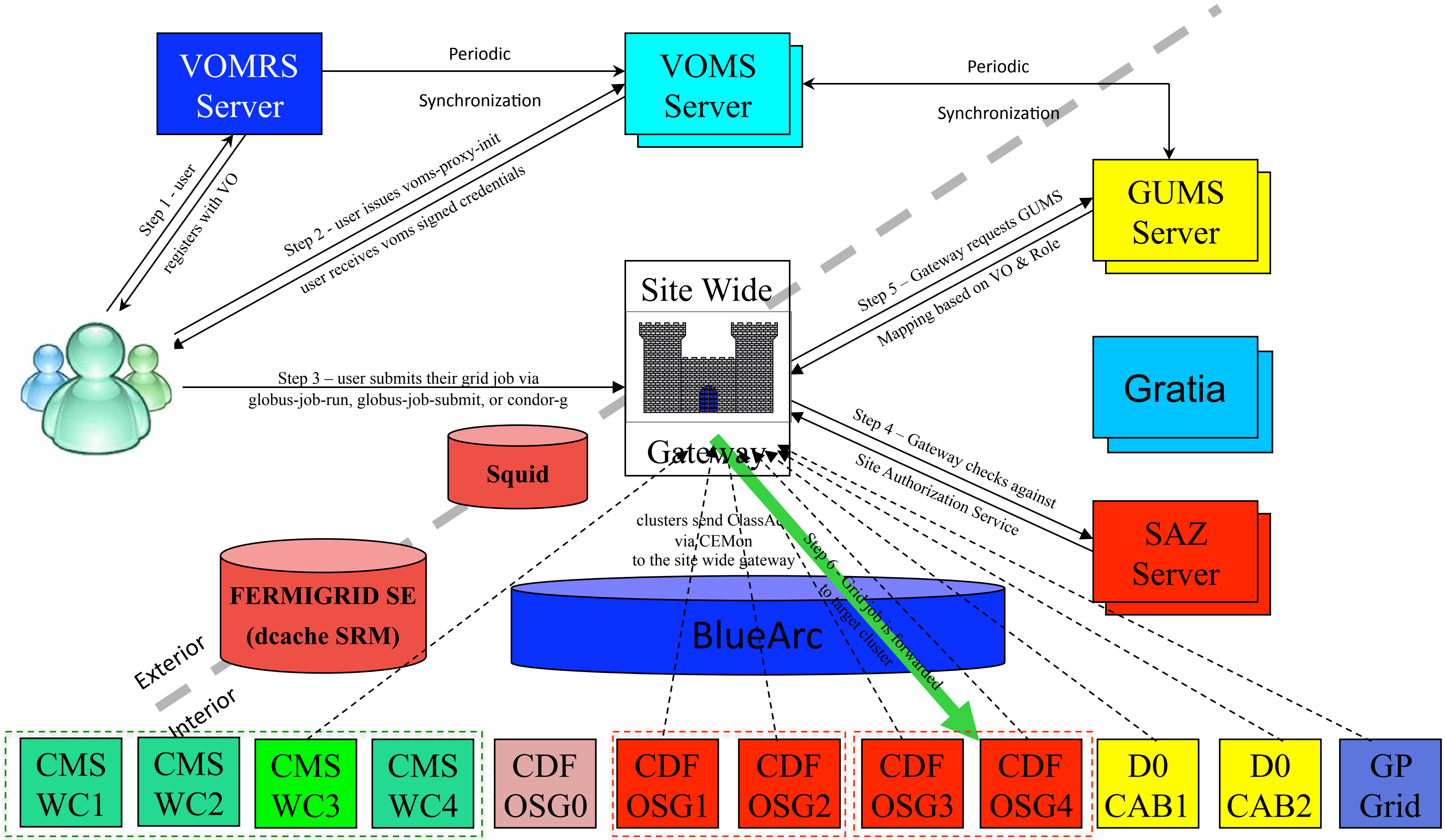
- Multiple interfaces and worker node configurations:
  - CDF – Kerberos + Condor
  - D0 – Kerberos + PBS
  - CMS – Grid + Condor
  - GP – Kerberos + FBSNG

# FermiGrid – Today

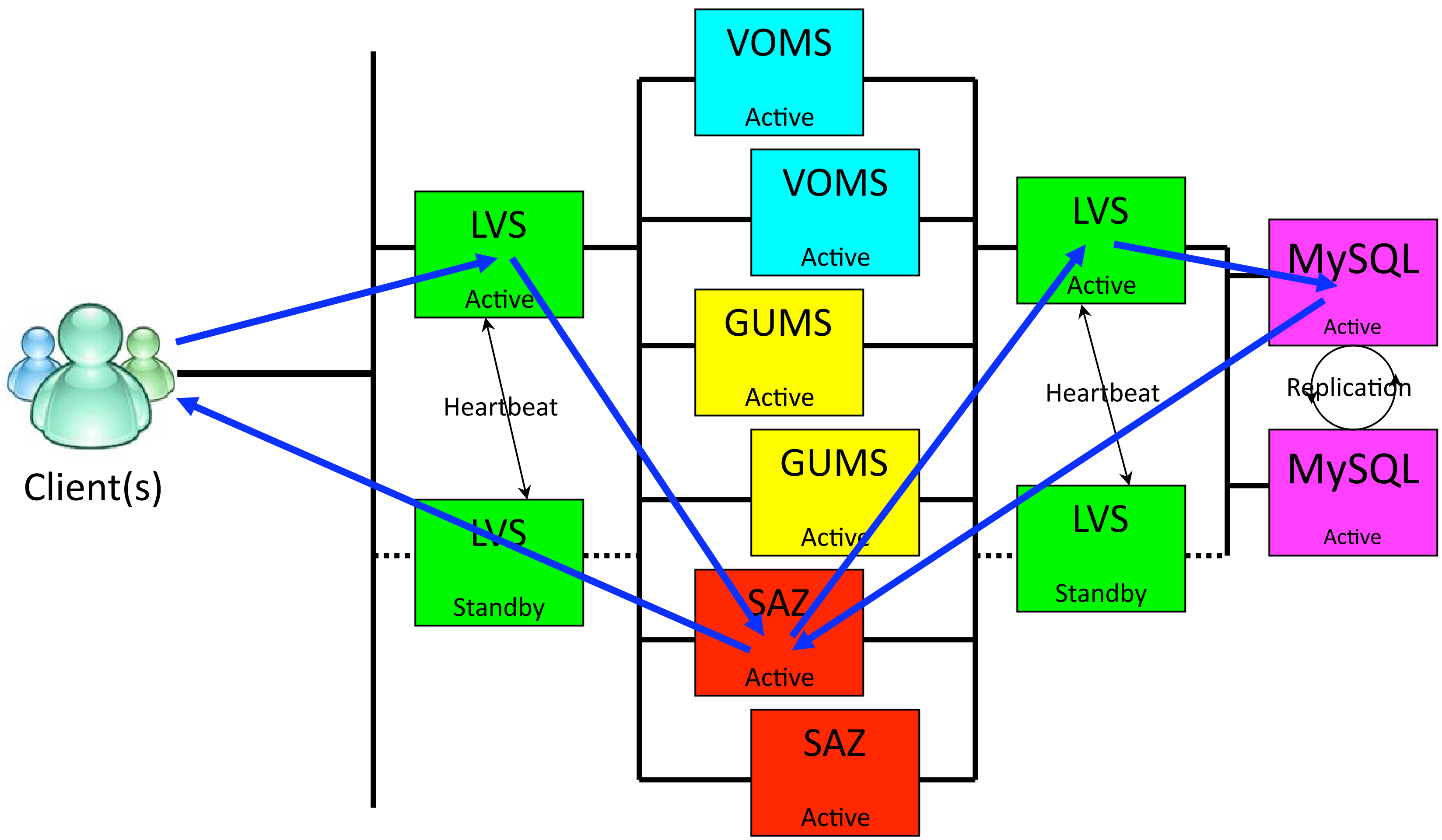
- Site Wide Globus Gatekeeper (FNAL_FERMIGRID).
- Centrally Managed Services (VOMS, GUMS, SAZ, MySQL, MyProxy, Squid, Accounting, etc.)
- Compute Resources are "owned" by various stakeholders:

| Compute Resources | # Clusters | # Gatekeepers | Batch System | # Batch Slots |
|---|---|---|---|---|
| CDF | 3 | 5 | Condor | 5685 |
| D0 | 2 | 2 | PBS | 5305 |
| CMS | 1 | 4 | Condor | 6904 |
| GP | 1 | 3 | Condor | 1901 |
| **Total** | **7** | **15** | **n/a** | **~19,000** |
| Sleeper Pool | 1 | 2 | Condor | ~14,200 |

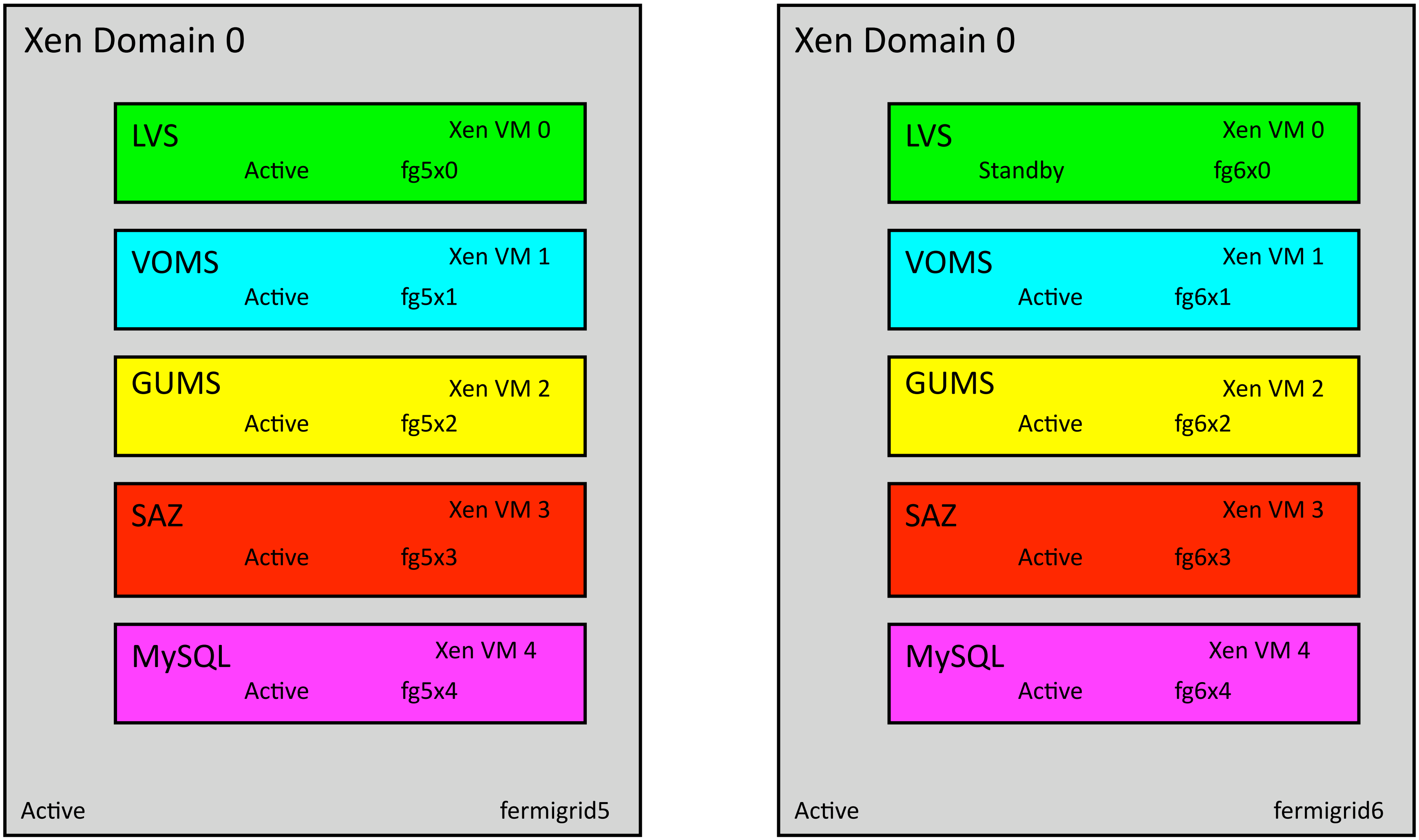
# FermiGrid - Architecture



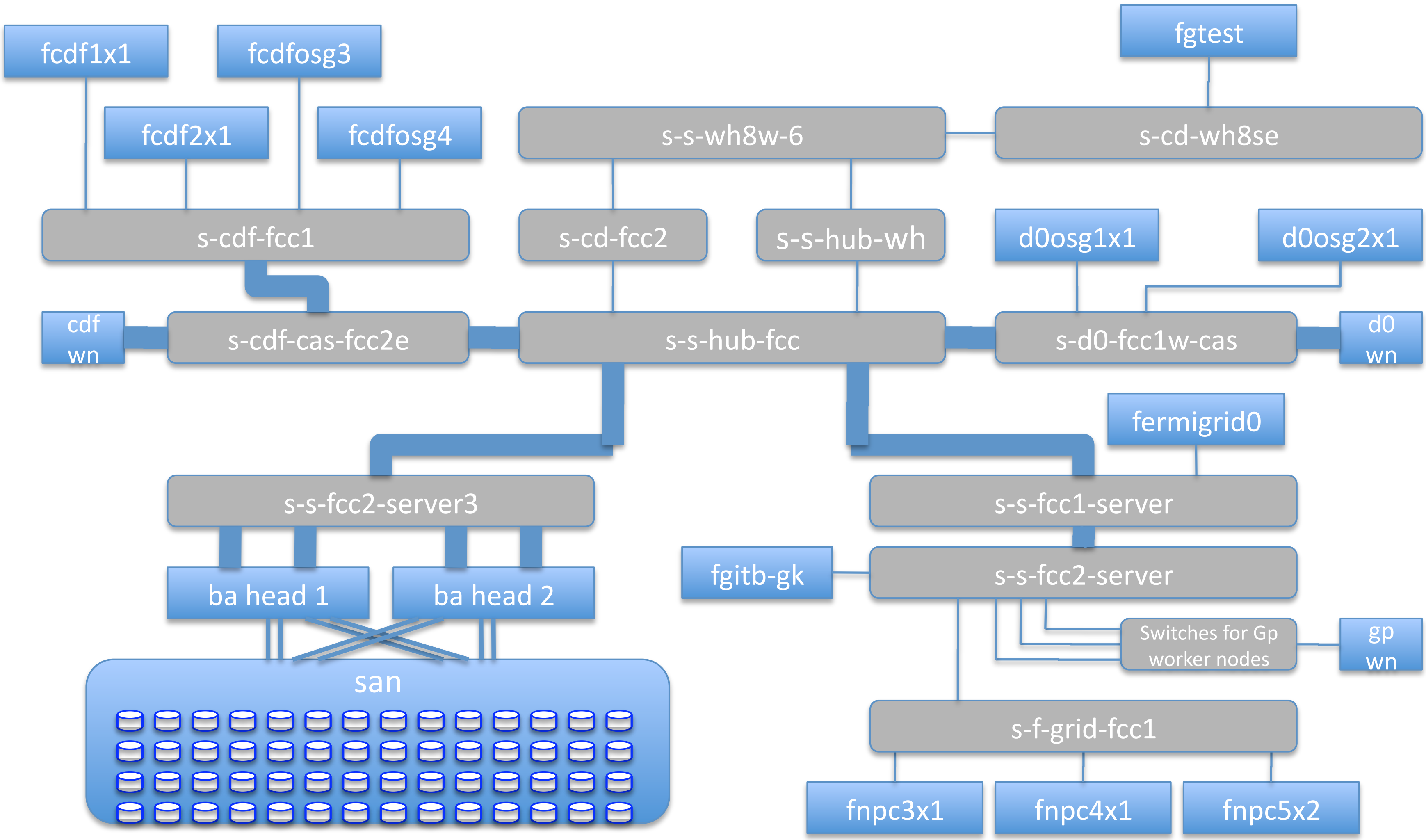
# FermiGrid HA Services - 1



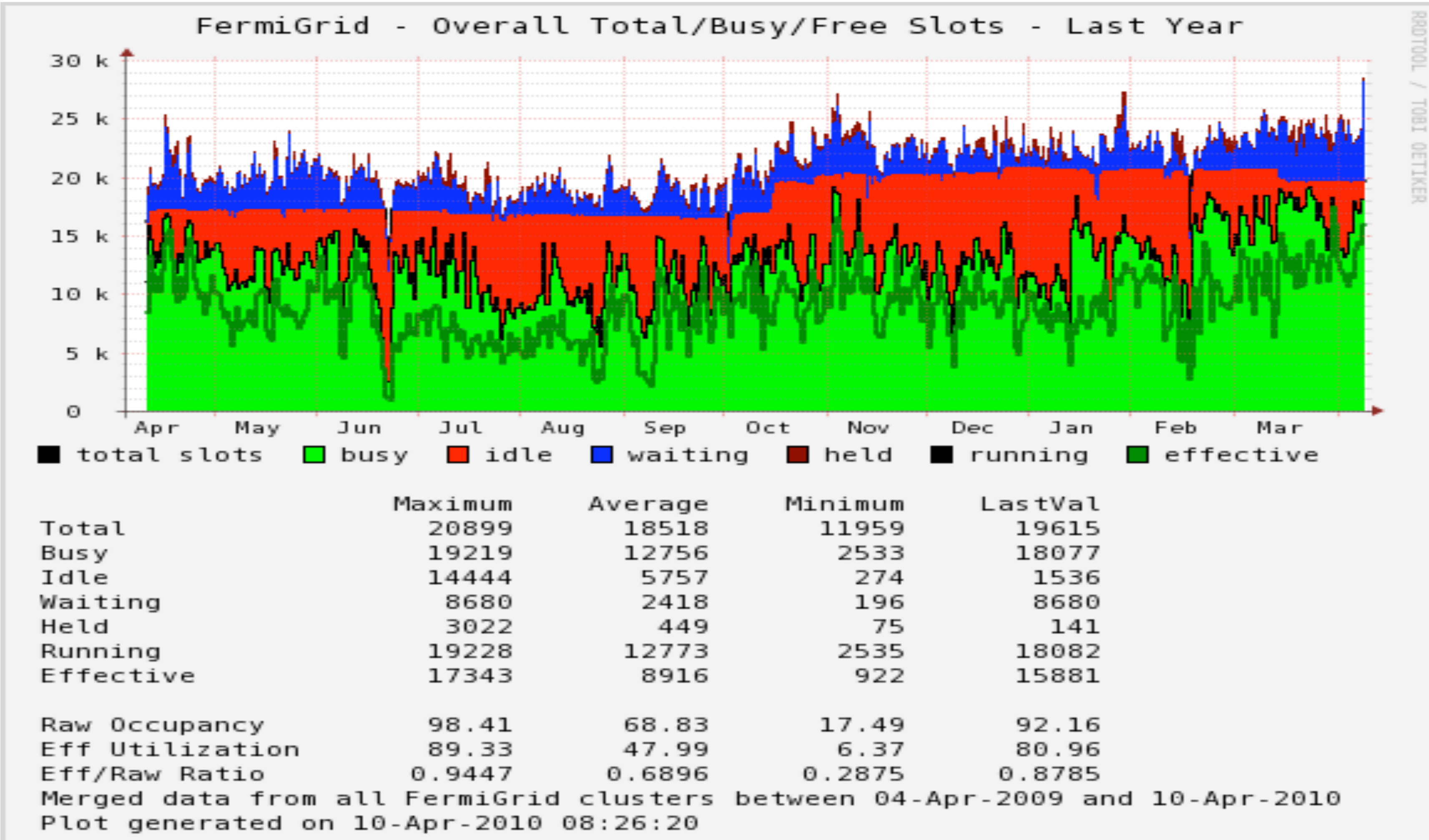
# FermiGrid HA Services - 2



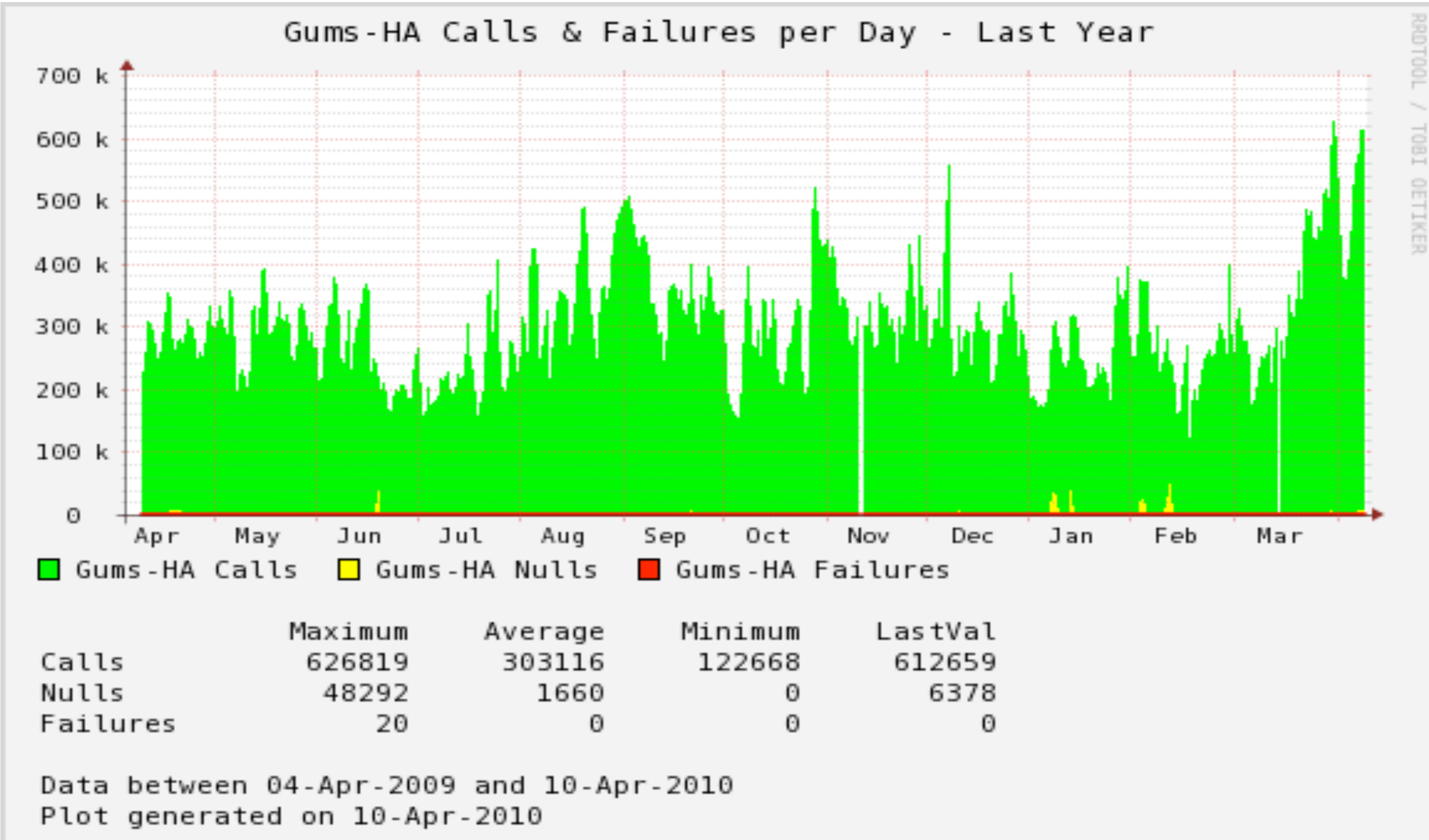
# (Simplified) FermiGrid Network



# FermiGrid Utilization

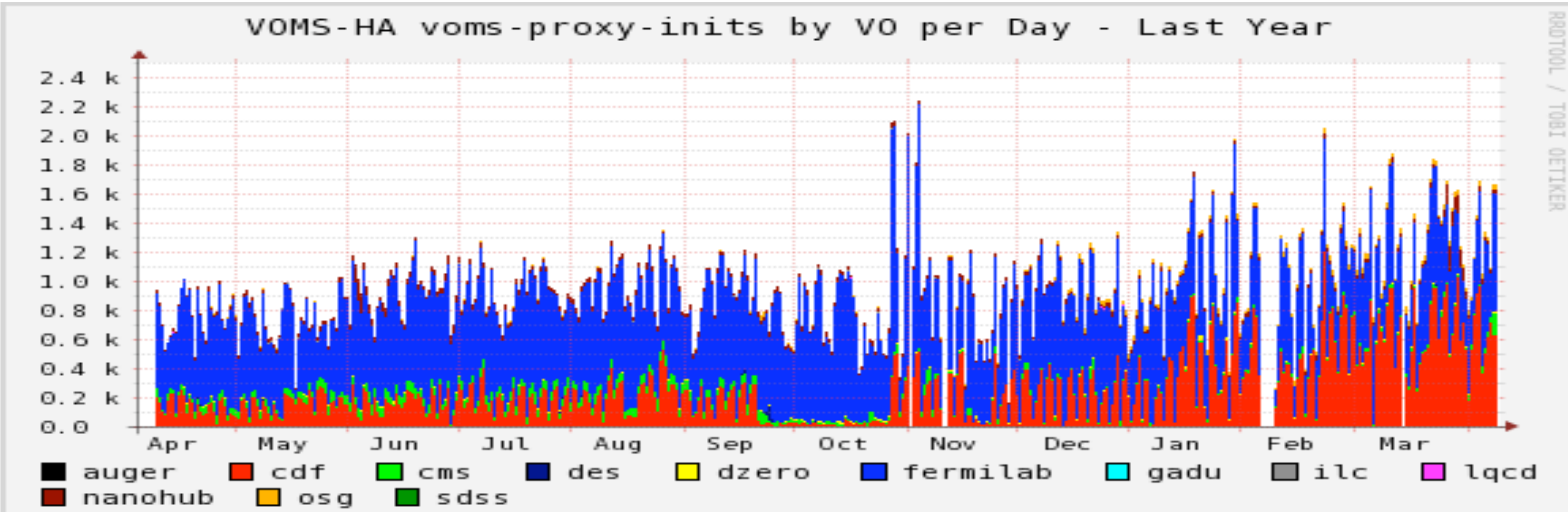


# GUMS calls



Gums-HA Calls & Failures per Day - Last Year
700 k
600 k
500 k
400 k
300 k
200 k
100 k
0
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
Jan
Feb
Mar
Gums-HA Calls
Gums-HA Nulls
Gums-HA Failures
RRDTOOL / TOBI OETIKER
Maximum Average Minimum LastVal
Calls 626819 303116 122668 612659
Nulls 48292 1660 0 6378
Failures 20 0 0 0
Data between 04-Apr-2009 and 10-Apr-2010
Plot generated on 10-Apr-2010

# VOMS-PROXY-INIT calls




| | Maximum | Average | Minimum | LastVal |
|---|---|---|---|---|
| auger | 0 | 0 | 0 | 0 |
| cdf | 1201 | 256 | 0 | 623 |
| cms | 155 | 32 | 0 | 155 |
| des | 101 | 1 | 0 | 0 |
| dzero | 29 | 4 | 0 | 7 |
| fermilab | 1689 | 619 | 0 | 813 |
| gadu | 0 | 0 | 0 | 0 |
| ilc | 8 | 1 | 0 | 0 |
| lqcd | 11 | 0 | 0 | 0 |
| nanohub | 174 | 24 | 0 | 30 |
| osg | 55 | 7 | 0 | 32 |
| sdss | 0 | 0 | 0 | 0 |

Data between 04-Apr-2009 and 10-Apr-2010
Plot generated on 10-Apr-2010

# Evolution

- You don't have to start with a massive project to transition to a Grid infrastructure overnight.

- FermiGrid was commissioned over roughly a 18 month interval:
  - Ongoing discussions with stakeholders,
  - Establish initial set of central services based on these discussions [VOMS, GUMS],
  - Work with each stakeholder to transition their cluster(s) to use Grid infrastructure,
  - Periodically review the set of central services and add additional services as necessary/appropriate [SAZ, MyProxy, Squid, etc.].

# Other Considerations

- You will likely want to tie your (?centrally managed?) administration/staff/faculty/student computer account data into your Campus Grid resources.
  - FermiGrid has implemented automated population of the "fermilab" virtual organization (VO) from our Central Name and Address Service (CNAS).
  - We can help with the architecture of your equivalent service if you decide to implement such a VO.

- If you have centrally provided services to multiple independent clusters [eg. GUMS, SAZ], you will eventually need to implement some sort of high availability service configuration.
  - Don't have to do this right off the bat, but it is useful to keep in mind when designing and implementing services.
  - FermiGrid has implemented highly available Grid services & we are willing to share our designs and configurations.

# What About Cloud Computing?

- Cloud Computing can be integrated into a Campus Grid infrastructure.

# Additional Resources

- FermiGrid
  - http://fermigrid.fnal.gov
  - http://cd-docdb.fnal.gov
- OSG Campus Grids Activity:
  - https://twiki.grid.iu.edu/bin/view/CampusGrids/WebHome
- OSG Campus Grids Workshop:
  - https://twiki.grid.iu.edu/bin/view/CampusGrids/WorkingMeetingFermilab
- ISGTW Article on Campus Grids:
  - http://www.isgtw.org/?pid=1002447

# Conclusions

- Campus Grids offer significant cost savings.
- Campus Grids do require a bit more infrastructure to establish and support.
  - This can be added incrementally.
- Many large higher education and research organizations have already deployed and are making effective use of Campus Grids.
- Campus Grids can be easily integrated into larger Grid organizations (such as the Open Science Grid or TeraGrid) to give your community access to larger or specialized resources.
  - Of course it's nice if you are also willing to make your unused resources available for opportunistic access.